

AMENDMENTS to the CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application.

Listing of Claims:

1. (currently amended) A computer-implemented method:
generating by a portion of a computing platform a set of cleaning attributes for each cleaned data record in a complete set of cleaned data records, said records each having a plurality of fields and being encoded by one or more computer storage drives or memory devices, said cleaning attributes indicating fields modified by a cleaning operation, wherein generating a set of cleaning attributes comprises performing an operation selected from a group comprising appending a set of cleaning attributes to each cleaned data record, prepending a set of cleaning attributes to each cleaned data record, distributing a set of cleaning attributes to each cleaned data record, and generating a cleaning attribute table, said computing platform having a hardware means for performing a logical process;
receiving by said portion of a computing platform a data feature identified within said cleaned data records for a subset of said complete set of cleaned data records;
determining by said portion of a computing platform a degree of correlation of said data feature to said indicated fields; and
responsive to said degree of correlation exceeding a threshold, identify identifying on a computing platform output device said data feature as having inaccurate data.
2. (previously presented) The method as set forth in Claim 1 wherein generating a set of cleaning attributes comprises generating a set of bit-mapped Boolean flags, wherein each Boolean flag corresponds to a field in a record.
3. (cancelled).
4. (previously presented) The method as set forth in Claim 1 wherein said data feature comprises a data feature selected from a group comprising a cluster, a trend, and a pattern.
5. (previously presented) The method as set forth in Claim 1 wherein generating a set of cleaning attributes comprises comparing each record in a raw data set to each record in a cleaned data set.

Claims 6 - 18 (cancelled)

19. (previously presented) A computer memory comprising:
a computer memory suitable for encoding software programs; and
one or more software programs encoded by said computer memory and configured to:
generate a set of cleaning attributes for each cleaned data record in a complete set of cleaned data records, said records each having a plurality of fields, said cleaning attributes indicating fields modified by a cleaning operation, wherein generating a set of cleaning attributes comprises performing an operation selected from a group comprising appending a set of cleaning attributes to each cleaned data record, prepending a set of cleaning attributes to each cleaned data record, distributing a set of cleaning attributes to each cleaned data record, and generating a cleaning attribute table;
receive a data feature identified within said cleaned data records for a subset of said complete set of cleaned data records;
determine a degree of correlation of said data feature to said indicated fields; and
responsive to said degree of correlation exceeding a threshold, identify said data feature as having inaccurate data.
20. (previously presented) The computer memory as set forth in Claim 19 wherein said software program configured to generate a set of cleaning attributes is further configured to generate a set of bit-mapped Boolean flags, wherein each Boolean flag corresponds to a field in a record.
21. (previously presented) The computer memory as set forth in Claim 19 wherein said data feature comprises a data feature selected from a group comprising a cluster, a trend, and a pattern.
22. (previously presented) The computer memory as set forth in Claim 19 wherein said software program configured to generate a set of cleaning attributes is further configured to compare each record in a raw data set to each record in a cleaned data set.

23. (previously presented) A system comprising:

- a computing platform having a hardware means to execute a logical process;
- an attribute generator portion of said computing platform configured to generate a set of cleaning attributes for each cleaned data record in a complete set of cleaned data records, said records each having a plurality of fields, said cleaning attributes indicating fields modified by a cleaning operation, wherein generating a set of cleaning attributes comprises performing an operation selected from a group comprising appending a set of cleaning attributes to each cleaned data record, prepending a set of cleaning attributes to each cleaned data record, distributing a set of cleaning attributes to each cleaned data record, and generating a cleaning attribute table;
- a data feature receiver portion of said computing platform configured to receive a data feature identified within said cleaned data records for a subset of said complete set of cleaned data records;
- a correlator portion of said computing platform configured to determine a degree of correlation of said data feature to said indicated fields; and
- an output portion of said computing platform configured to, responsive to said degree of correlation exceeding a threshold, identify said data feature as having inaccurate data.

24. (previously presented) The system as set forth in Claim 23 wherein said attribute generator is further configured to generate a set of bit-mapped Boolean flags, wherein each Boolean flag corresponds to a field in a record.

25. (previously presented) The system as set forth in Claim 23 wherein said data feature comprises a data feature selected from a group comprising a cluster, a trend, and a pattern.

26. (previously presented) The system as set forth in Claim 23 wherein said attribute generator is further configured to compare each record in a raw data set to each record in a cleaned data set.